FOODBORNE SOURCES OF BACTERIA ASSOCIATED WITH HUMAN OBESITY

A Senior Scholars Thesis

by

KATHERINE GRACE MCELHANY

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

April 2008

Major: Biology

FOODBORNE SOURCES OF BACTERIA ASSOCIATED WITH HUMAN OBESITY

A Senior Scholars Thesis

by

KATHERINE GRACE MCELHANY

Submitted to the Office of Undergraduate Research
Texas A&M University
in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by:

Research Advisor: Suresh Pillai
Associate Dean for Undergraduate Research: Robert C. Webb

April 2008

Major: Biology

ABSTRACT

Foodborne Sources of Bacteria Associated with Human Obesity (April 2008)

Katherine Grace McElhany
Department of Biology
Texas A&M University

Research Advisor: Dr. Suresh Pillai
Department of Poultry Science & Nutrition and Food Science

Recently published research has suggested that the microbial ecology of the digestive system may play a role in obesity. Obese people have been shown to have a higher proportion of bacteria from the Firmicutes division and a lower proportion of bacteria from the Bacteroidetes division in their gut. The goal of this study was to characterize the microbial communities in specific foods using a combination of microbiology and metagenomic techniques. The ultimate goal of this project was to identify specific foods that may be introducing the Firmicutes and Bacteroidetes bacterial groups into the human gut. Nine commonly consumed foods, both natural and processed, were selected for this study and were purchased from a retail outlet in College Station, TX. These included wheat bread, whole milk, spinach, low-fat yogurt, medium cheddar cheese, 80/20 ground beef, salmon, banana, and skinless chicken breast samples. The food samples were plated on Tryptic Soy Agar (TSA) media to determine the aerobic and anaerobic bacterial loads and to isolate bacteria. The total microbial community was extracted from these food samples and the 16S rRNA gene sequences of the microbiome

were PCR amplified. The PCR amplicons were sequenced using pyrosequencing and the metagenomic data was analyzed using bioinformatic approaches. The culture-based data suggests that there is a substantial bacterial load on some of the food items, ranging from as low as $5.9x10^2$ CFU/gm to $2.8x10^6$ CFU/gm. Results of the pyrosequencing data indicate that the cheese, ground beef, salmon, milk, and chicken breast contain significant amounts of Bacteroidetes and/or Firmicutes. The results of this study suggest that 1) foods harbor a variety of microbial populations including those that have been associated with human obesity and 2) the consumption of specific food types could be influencing the types of microorganisms inhabiting the human gut.

DEDICATION

This thesis is dedicated to my parents, who have always expected the best of me.

ACKNOWLEDGMENTS

I would like to thank especially my advisor Dr. Suresh Pillai, who has been the epitome of support and encouragement this year. In addition, I wish to thank Martha Cepeda, Dr. Palmy Jesudhasan, and the rest of the lab for their endless patience and help. I extend my sincerest gratitude to Dr. Scot Dowd at the USDA for all his help with the pyrosequencing. I would also like to thank the Office of Undergraduate Research for the opportunity to gain this type of research experience at the undergraduate level.

NOMENCLATURE

CFU	Colony Forming Unit
E-value	Expectation Value
FAME	Fatty Acid Methyl Ester
PCR	Polymerase Chain Reaction
SIM	Similarity Index
TSA	Tryptic Soy Agar
USDA	United States Department of Agriculture

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Most of the bacterial organisms that inhabit the human body live in the gastrointestinal tract. The ecosystem of the human intestine is incredibly diverse, containing between 10^{13} and 10^{14} microorganisms (6). New research is just beginning to prove how much the bacterial inhabitants of the intestine contribute to digestion and the overall health of the body. Intestinal microflora are integral parts of many essential processes, but perhaps most importantly the extraction of energy from foods in the form of polysaccharides. Select microbial flora in the human gastrointestinal tract produce enzymes that allow them to turn normally indigestible polysaccharides into calories. These enzymes are not produced in humans, meaning that this process would not occur in the absence of gut microbiota (14).

In the past, research has focused almost solely on bacterial species that can have a detrimental effect on human health. This is understandable, as these bacteria are occasionally dangerous, often contagious, and always inconvenient. However, this sole focus on has led to an erroneous scientific attitude that bacteria fall into one of two categories: dangerous or inconsequential. Very little is known about the billions of

---

This thesis follows the style of Applied and Environmental Microbiology.

bacteria that naturally inhabit the human body and new research is challenging the assumption that these inhabitants do not have an impact on health. On the contrary, some bacteria are demonstrating the potential to be very beneficial.

The idea of beneficial bacteria, commonly referred to as probiotics, is not a new idea, but it is one that has only become prominent in the last decade. Products have been marketed by multiple food companies that claim to use beneficial bacterial species to improve digestion and immunity. The fact that these products remain on grocery store shelves indicates that they have been accepted by consumers. Notable examples are Dannon Activia®, Dannon DanActive®, and Yoplait Yo-Plus®. Dannon Activia® and Yoplait Yo-Plus® are marketed to improve digestive function and contain different cultures of *Bifidobacterium* (Product Packaging). Dannon DanActive® contains cultures of *Lactococcus casei* and is marketed to strengthen the immune system (Product Packaging). The successful marketing of these products is important for two reasons. First, the fact that consumers will purchase these products shows that the general public has overcome at least some of the negative imagery associated with "bacteria" and has accepted the idea that some bacteria may be beneficial. This may also be due to the fact that the marketing and packaging for these products does not refer to the product inhabitants as "bacteria", but rather as "cultures". Secondly, it demonstrates that bacteria consumed in the form of food can have a noticeable, if temporary, affect on health.

Researchers are studying how gut bacteria extract energy from foods because the process is poorly understood and these processes are thought to strongly influence overall caloric intake. It is also hoped that more knowledge could help to identify the various factors involved in weight gain and loss, especially in regard to obesity (3). Obesity is one of the most profound and immediate threats to modern public health. In 2000, it was estimated that approximately 20% of the population of the United States was obese and this statistic is predicted to rise. This reality is not unique to the United States—many regions of the world including Latin America, Europe, Southeast Asia, and the Middle East have seen a remarkable rise in obesity among their citizens (13). Complications of obesity can include, but are not limited to, hypertension, diabetes, cardiovascular disease, stroke, respiratory complications, osteoarthritis, and several varieties of cancer (3, 13). Health problems caused by obesity place an extreme financial burden on both patients and health-care systems.

For years, the cause of obesity has been thought of simply in terms of calories expended versus calories consumed. More recent research, however, has shown that host genetics and the physiology of gut-associated bacteria are perhaps some of the more important factors in an individual's predisposition to weight gain or loss (4, 13). Intestinal bacteria that process foods more efficiently would produce more calories. These excess calories would then lead to increased weight gain, regardless of consumption. In short, two people with the exact same diet could absorb different numbers of calories depending on the types of microbiota inhabiting their gastrointestinal tract. Some people could simply

be better environments for these “efficient” gut microbiota, predisposing them to weight gain (4, 14). It is hoped that future research to understand the role of genetics in obesity will have some potential therapeutic value for morbidly obese patients.

Ley *et al*. was the first to show that obese mice had a higher proportion of Firmicutes and a lower proportion of Bacteroidetes inhabiting their gut than comparable lean mice. (14). Bacteroidetes and Firmicutes are two of the major divisions of bacterial organisms. Both groups are generally considered benign and are commonly found in the natural world (4). Research published by Ley *et al*. also showed the same basic trend in humans, with obese individuals demonstrating an increased proportion of Firmicutes and a decreased proportion of Bacteroidetes in their gut. Interestingly, the research showed that the proportion of Bacteroidetes increased as the subjects lost weight (15).

The research performed by Ley *et al*. also demonstrated that there was a strong link between kinship and gut microbiota, meaning that related mice had much more similar gut microbiota than those that were not related. Mothers and offspring displayed very similar intestinal flora, as did siblings (14). These findings support the theory that a predisposition to colonization by certain bacteria is an inheritable factor. In 2004, a similar group of researchers published a paper that revealed how they had transplanted a microbial community from the intestine of normal mice into specially designed “germ-free” mice. After the transplant, the recipient mice experienced a rapid weight gain without any increase in food consumption (26). The results of this experiment make a

convincing argument that the microbial inhabitants of our intestines play an essential role in harvesting energy and nutrients from foods. The results also suggest that, contrary to popular belief, caloric intake may not be the only variable involved in weight gain or loss.

In the past few years, several research papers have been published that extensively detail the biodiversity of the human intestine (5, 6, 8, 24). In each of these experiments, the end result was accomplished by using a variety of molecular methods that are slowly shedding light on the molecular diversity of the gut-associated microbial populations. These relatively new methods employ 16S rRNA to collectively characterize and identify entire communities of bacteria. Culture-based methods of identification have severe limitations, especially when a large group of microorganisms are being analyzed. Anaerobic bacteria are also considerably more difficult to culture than aerobic organisms, but the majority of microorganisms inhabiting our digestive tract are anaerobic. The use of ribosomal RNA sequences (rRNA sequences), however, has allowed researchers to gain a much better understanding of the extraordinary bacterial diversity that inhabits the human digestive tract. This development is significant because it allows scientists to analyze microbes as a part of a larger community—rather than simply on an individual basis.

Human fetuses in the womb are completely sterile. The bacterial colonization of the infant occurs as a result of exposure to the environment and either breast or bottle

feeding. It has been shown that the microbial inhabitants of the infant's intestines are determined by the mother's breast milk and that the diversity of this community shifts over time as the child progresses to other foods (19). It is therefore obvious that the foods we consume at the beginning of our lives directly impact the species of bacteria that colonize our digestive tract. Building on this information, one question of scientific interest that presents itself is whether or not the foods consumed throughout the remainder of our lives could continue to influence changes in gut microbiota. Therefore, even though recent research has greatly improved our understanding of the intestinal microflora endogenous to humans, little information is known about the diversity of the microbial populations present in our foods and whether they have an impact on the microbial diversity of the gastrointestinal tract of individuals.

This objective of this project was to identify whether microorganisms that inhabit different types of food could be linked to key microbial divisions that have been recently linked to obesity. As was previously mentioned, recently published papers have linked obesity to the proportion of two major microbial divisions in the human digestive tract, Bacteroidetes and Firmicutes. The primary objective of this study was to test selected foods for the presence of these organisms. Nine different foods were chosen for analysis of bacterial load: wheat bread, whole milk, spinach, low-fat yogurt, medium cheddar cheese, 80/20 ground beef, salmon, banana, and boneless, skinless chicken breast. These foods were chosen based upon their prevalence in the American diet. The underlying

hypothesis was that a better understanding of the microbial communities inhabiting commonly consumed foods would help to clarify the ways in which humans may be exposed to these bacteria and suggest whether the consumption of certain foods could predispose a person to obesity.

CHAPTER II

METHODS

Preparation of food products

Purchase of food products

All food products were purchased at a local chain grocery store in College Station, Texas. Each purchased food product was verified to ensure that the item was not past its expiry date. The nutritional information for each of the purchased food products was recorded if this data was available on the product packaging.

Homogenization of food products

Each food product was homogenized to ensure an accurate sample for plating and DNA extraction. For aerobic and anaerobic plate counts to be accurate, bacteria present in the food must be evenly distributed throughout the sample. The homogenization also served to ensure that any bacteria present in the interior of the food product would included in the subsequent samples.

Homogenization was achieved by placing 50 g of food product into a sample bag with a filter insert (VWR, West Chester, Penn.). If necessary, the sample was mixed or shaken before the 50 gram sample was taken. The filter bag was filled with 450 mL of sterile Butterfield's phosphate buffer dilution water. The filter bag was then placed into a

Seward Stomacher 400 (Brinkmann, Westbury, N. Y.) and processed for 120 seconds on the "fast" setting. This procedure was repeated three times for each food product, resulting in three different homogenized samples. Because the milk samples were already in a liquid and homogenous state, they were not mixed with the Butterfield's solution or stomached.

Aerobic and anaerobic plate counts

Aerobic plate counts

A 0.1 mL aliquot was removed from the filtered section of each sample bag and used to make a set of 10-fold serial dilutions in Butterfield's dilution buffer. The serial dilutions, along with 0.1 mL of the undiluted sample from the sample bag, were plated on Tryptic Soy Agar, resulting in plated dilutions of 10^{-1}, 10^{-2}, 10^{-3}, and 10^{-4}. A dilution of 10^{-5} was plated for some food products if a high bacterial load was anticipated. The final result of this procedure was three separate sets of plated food product, each originating from one of the three homogenized food samples and each set containing four or five plates. The plates were incubated for four days at 27°C. After four days of incubation, the plates were removed and colony growth for each plate was assessed and counted. Based upon morphology, certain bacterial colonies were isolated and sub-cultured for subsequent identification.

Anaerobic plate counts

For comparison purposes, only those foods that showed a bacterial load from aerobic plating were subjected to anaerobic plate counts. The protocol for the anaerobic plate counts was nearly identical to the procedure listed previously for aerobic plate counts. The only difference between the two procedures was that after the dilutions for each sample were plated onto TSA, the plates were placed into a GasPak EZ Anaerobe Gas Generating Pouch System with Indicator (BD, Franklin Lakes, N. J.). As suggested by the manufacturer, a paper towel soaked in 5 mL of distilled water was added to each pouch to enhance the anaerobic conditions to facilitate bacterial growth. Each pouch contained a maximum of four plates and was sealed tightly. Plates were checked daily during incubation to ensure that the indicators showed anaerobic conditions for each pouch. After an incubation period of five days at 27°C, the plates were removed from the Anaerobe bags and colony growth was assessed and counted. Selected colonies were isolated and sub-cultured based on morphology.

Colony subculturing and isolation

Colonies selected for sub-culturing were streaked for isolation on TSA plates using standard procedures. For colonies taken from anaerobically incubated plates, the streaked TSA plates were immediately placed into GasPak pouches according to the protocol described in the previous section. To minimize loss of anaerobic bacteria due to oxygen exposure, the contents of each anaerobic pouch were counted, streaked for isolation, and the sub-cultured plates returned to anaerobic pouches within 15 minutes of

removal from anaerobic conditions. Sub-cultured plates were incubated for 4-5 days at 27°C or until significant colony growth was observed.

An isolated colony from each subculture was then inoculated in a tube of Tryptic Soy Broth and placed in a shaking incubator at 27°C. Colonies taken from anaerobically incubated plates were inoculated in TSB tubes and immediately placed back into anaerobic pouches. Once again, this procedure was completed within 15 minutes to minimize oxygen exposure to sensitive anaerobic species. Each inoculated culture was incubated until turbid—generally around 24 hours, but somewhat variable among samples. For every grown aerobic and anaerobic sample, 1.5 mL of bacterial culture was placed into each of three labeled cryovials and centrifuged for 1 minute at 15,000 x g. The supernatant of each vial was discarded and the bacterial pellet resuspended in a glycerol/TSB solution. Each sample vial was then immediately stored in a freezer at -80°C until further use.

Identification of aerobic culture isolates

A pure culture of each aerobic isolate was plated on Tryptic Soy Agar and delivered to the Texas Plant Disease Diagnostic Laboratory (TPDDL) at Texas A&M. Unfortunately, TPDDL did not have the capacity to analyze samples isolated from the anaerobic plates. The isolates were analyzed by the Fatty Acid Methyl Ester Procedure. This method consists of growing each isolate on a specific media at 28±1ºC for 24±1 hours. Approximately 40 mg of fresh weight cells was harvested from the plates and subjected

to whole cell fatty acid extraction and esterification. The bacterial methyl ester extract was then placed in a specialized vial and analyzed by gas chromatography. A computer recorded the specific fatty acid profile for each isolate, which was then compared to a database of information. This database contained profiles for all known major plant pathogens, but also has access to profiles for many other bacterial groups as well. By comparing the results to profiles of known bacteria, the analysis program developed a Similarity Index (SIM) for each isolate based on at least 10 different comparisons. The identified bacterial species with the highest SIM is generally the closest match to the analyzed isolate. A SIM of greater than 5 is recognized as a good match on the genus and species level and a SIM of greater than 2 is seen as a good match on the genus level. Once the analysis results were received from the TPDDL, isolates with a SIM greater than 2 were recorded as their highest SIM match. Isolates with a SIM less than 2 were recorded as "Inconclusive". Since this research project concerns higher bacterial phylums instead of individual species, only the genus of each identified isolate was recorded.

Molecular methods

DNA extraction from food products

Extraction of community DNA was only performed on those foods that demonstrated bacterial load from aerobic plating. A 125 mL volume of the filtered food homogenate was removed from the stomacher sample bag and placed into a sterile centrifuge bottle.

The sample was centrifuged at 10,000 x g for 10 minutes. After completion, the supernatant was discarded and another 125 mL of filtered food homogenate was added to the centrifuge bottle. The bottle was vortexed to resuspend the pellet and placed back in the centrifuge for 10 minutes. The sample supernatant was discarded, the pellet washed in 10 mL of Butterfield's, and the sample centrifuged again for 10 minutes. The resulting supernatant was discarded and the pellet resuspended in 1 mL of Butterfield's. This procedure was altered slightly for milk, which is already in a liquid state. Instead of being mixed with Butterfield's and stomached, the milk was homogenized by shaking and then added directly to the centrifugation bottle.

DNA was extracted from the food samples using the Maximum Yield Protocol of the Ultra Clean Soil DNA Kit (MoBio, Carlsbad, Calif.). Extracted DNA was analyzed for yield and purity using a NanoDrop spectrophotometer and stored at -20°C for future applications. Results of spectrophotometer analysis are shown in in Appendix A.

Pyrosequencing of food sample community DNA

For analysis of metagenomic content by the pyrosequencing method, samples of extracted community DNA were sent to the USDA Livestock Issue Research Center located in Lubbock, Texas. For each food sample, the extracted community DNA sample with the highest concentration of DNA nucleic acids and purity as shown by spectrophotometer analysis was selected for pyrosequencing. A detailed explanation of the principle and methods involved in the pyrosequencing procedure is included in

Chapter III. Analysis of the samples was completed using the bTEFAP FLX pyrosequencing procedure as detailed below.

A 50 µl PCR reaction was performed for each food sample using 1µl of extracted DNA and a HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, Calif.). The PCR reaction was performed using the 16S universal bacterial primers 530F (5'-GTG CCA GCM GCN CGC G) and 1100R (5'-GGG TTN CGN TCG TTG). These primers were designed to selectively amplify a 600bp variable region of 16rRNA gene segments. The PCR conditions were as follows: 94ºC for 3 minutes, followed by 32 cycles of 94ºC for 30 seconds, 60ºC for 40 seconds, and 72ºC for 1 minute; and concluded by 72ºC for 5 minutes. After completion of the first PCR run, a second PCR was performed using specifically designed fusion primers designated as LinkerA-Tags-530F and LinkerB-1100R. All other conditions for the secondary PCR were the same as the first reaction. This second PCR was important for the elimination of bias that otherwise may occur during initial amplification of the template.

The final amplified PCR products were purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, Beverly, Mass.). The purified DNA template strands were measured for size and concentration using a Bio-Rad Experion Automated Electrophoresis Station (Bio-Red Laboratories, Hercules, Calif.) and TBS-380 Fluorometer (Turner Biosystems, Sunnyvale, Calif.). The concentration of the DNA was adjusted to $9.6x10^6$ double-stranded DNA molecules/µl. These DNA sequences had an

average length of 625bp. After combination with approximately 9.6 million DNA capture bead, the DNA template strands were amplified using emulsion PCR. Bead-attached double-stranded DNA molecules were separated into single-straned molecules using NaOH. After attachment of necessary sequencing primers, the samples were run on a 70x75 GS PicoTiterPlate using a Genome Sequencer FLX System (Roche, Nutley, N. J.). In addition to the 6 food product samples, 94 additional samples associated with unrelated projects were included in the same sequencing run. All procedures relating to FLX sequencing were completed using Genome Sequencer FLX System manufacturers instructions (Roche, Nutley, N. J.)

A program was written in C# for the purpose of bTEFAP tag design. All combinations of 10-mer oligonucleotide tags with a GC% between 40% and 60% were generated using this program and 20 separate tags chosen from the results. Programs written in C# were also used for the analysis involved in the processing of the pyrosequencing data. This software was developed within a Microsoft® .NET (Microsoft Corp., Seattle, Wash.) environment. Reads taken from FLX sequencing output files were trimmed for quality and exported into a multi-FASTA file, from which tags were extracted and sorted into designated files based on the tag sequence. All tags with less than 100% homology to the given sample designation were discarded, as were sequences consisting of less than 150bp after trimming. The FASTA files were then assembled using the CAP3 program and processed to generate a cumulative file that contained tentative consensus sequences for each sample, as well as the number of sequences included in each consensus (9). The

secondary FASTA file for each sample was then compared against a custom database derived from the RDP-II database using the BLASTn program (2, 17). Finally, the resulting data for each food sample was sorted by genera and subjected to phylogenetic analysis.

The pyrosequencing output data contained bacteria identified to the species level, but this level of identification was not necessary for the purposes of this project. All results for each food product were therefore sorted by genera. Sequences with an Expectation Value (E-value) higher than 10e-80 were not included in the final data analysis because of their low correlation to the 16S rRNA gene sequences of known organisms. For each food product, the adjusted data was sorted by phylum using the National Center for Biotechnology Information Taxonomy Database (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi/) and the proportions of different phyla assessed. Additionally, the pyrosequencing data was analyzed for the presence of known human gut microbiota established by scientific literature and the phylogenetic analysis repeated including only these organisms (5, 7, 8, 23, 24). Lastly, regression analysis was performed to examine the relationship between the proportion of Bacteroidetes or Firmicutes and the nutritional content of the food products. Overall percentage of Bacteroidetes or Firmicutes was correlated with the calories per gram, fat grams per gram, or total carbohydrates per gram for each food product to determine if a relationship existed.

CHAPTER III

PYROSEQUENCING

Introduction

Pyrosequencing is a relatively new molecular technique with an incredible potential for metagenomic analysis. It is based upon what is known as a "sequencing-by-synthesis" method, utilizing specific enzymes to record each nucleotide inserted into a complementary DNA strand (1). The pyrosequencing technique has been used successfully to evaluate the microbial diversity of soil samples, detect medically significant pathogens, and distinguish different species of *Mycobacteria* (12, 20, 25). Although the concept behind pyrosequencing was developed in the 1980s, the actual procedure was not presented until the mid-1990s by a group of researchers at the Royal Institute of Technology in Stockholm (1, 22).

The bacterial 16S rRNA genes are highly variable genetic regions that are commonly used for taxonomic analysis. These variable regions are essentially fingerprints that, when compared against a 16S database, can determine the identity of an organism and its evolutionary relationship to other organisms. The highly variable 16S regions are surrounded by highly conserved segments that allow amplification of the region using specific primers (12). The newest machine developed by 454 Life Science™, the GS FLX sequencer, produces sequences between 200-300 nucleotides. However, researchers have determined that, if the pyrosquencing procedure and analysis are performed

correctly, fragments of 16S rDNA as short as 100 nucleotides can be just as effective for making taxonomic designations as long 16S rDNA sequences. (16).

Principle of pyrosquencing procedure

The pyrosequencing principle is based on the concerted action of four separate enzymes: DNA polymerase, sulfurylase, luciferase, and apyrase. The polymerase enzyme moves along the template DNA strand, incorporating nucleotides into the growing complementary DNA strand. Each time a nucleotide is added, pyrophosphate molecules are released into the surrounding environment. These pyrophosphates act as substrates for the sulfurylase enzyme, which converts them into ATP molecules. The ATP then reacts with the luciferase enzyme to produce a light reaction. The machine is able to read the light reaction and determine which nucleotide was incorporated into the sequence. The apyrase enzyme is required to degrade pyrophosphate and ATP at the end of each cycle to prevent excess products from interfering with polymerase activity (1, 21, 22). The sequencing results for each analyzed DNA fragment are recorded in the form of a pyrogram, which is a chart consisting of a series of peaks. Each peak represents a nucleotide addition and the slope and level of the peaks provide information regarding the activity of the enzymes (21).

Several preparatory steps are required before the template DNA can be subjected to analysis by pyrosquencing. The first step is PCR of the DNA sample, usually in oil emulsion, with universal bacterial primers. The PCR-amplified DNA template must then

be purified to eliminate primers and substrates used the PCR reaction (21). Immobilization of the DNA template is also required for pyrosequencing. This is accomplished by attachment of the amplified sequences to streptavidin-coated magnetic beads. To be sequenced correctly, a bead must contain only one DNA sequence. (10, 21). The bead-attached templates are then deposited on a specially designed plate that allows determination of the enzymatically produced light reactions. Past experimentation has optimized the level of enzymes needed in the pyrosequencing reaction for maximum sequencing efficiency to occur. Automated machines now add these necessary elements automatically during the sequencing run (21). Nucleotides are added automatically at select intervals throughout the sequencing run by microfluidics.

After completion of the sequencing run, results are subjected to analysis by a series of specifically-designed programs (25). These programs contain algorithms that check the sequences for quality results, rejecting sequences that contain a certain percentage of ambiguous nucleotides. Research has established that an important part of maintaining accurate pyrosequencing results is eliminating poor quality sequences (10). Other programs trim the sequences for quality and subject them to analysis via the appropriate database.

Advantages and disadvantages

The pyrosequencing procedure has many significant advantages. All sequencing is accomplished simultaneously, which allows the process to be completed much faster.

After preparation of the template DNA is complete, over a hundred thousand sequences can be generated from a single sample in a matter of hours. Pyroquencing also eliminates many of the time and cost constraints involved in the cloning and sequencing of individual 16S rDNA segments (10). Perhaps most importantly, the pyrosequencing procedure is unique in that monitoring of the enzymatic light reaction permits real-time analysis of the sequencing procedure—something that other sequencing methods have yet to provide (22). One evaluation of accuracy revealed that levels of error associated with pyrosequencing were less or comparable to the levels of error established for other DNA sequencing methods. Substitution errors are also less likely to occur in pyrosequencing than in other DNA sequencing methods (10). Although it has been shown to be possible and accurate, pyrosequencing is not yet the ideal method for whole-genome sequencing because of sequence length limitations (21).

One of the major disadvantages associated with pyrosequencing is cost. Although the average cost-per-sequence is lower than other methods, the cost required for to maintain and run the sequencer can be prohibitive. However, as this technique is improved, the entire process will likely become more affordable. The pyrosequencing procedure can also encounter problems when sequencing homopolymeric regions and polymorphic regions, which can result in the production of ambiguous light reactions that are difficult to decipher (10, 12, 21). An error resulting in loss of synchronization with other DNA sequences can also result in an indeterminate pyrogram for that fragment. Investigation

has shown that some poor quality results may be the result of two or more different sequences attached to the same immobilization bead (10).

Most microbial species found in nature have yet to be cultured in the laboratory (16). This fact makes effective and efficient DNA sequencing one of the essential molecular technologies required for the future of scientific research. The pyrosequencing method has demonstrated potential in common DNA sequencing objectives such as single-nucleotide polymorphisms, tag sequencing, metagenomic and taxonomic analysis, and the determination of the secondary structure of some proteins (21). The technique has also undergone preliminary evaluation for clinical use (12, 25). Pyrosequencing is currently the fastest available option for the sequencing of PCR products and some have predicted that pyrosequencing represents the future of DNA sequencing (16, 21). As technology advances, pyrosequencing automation, efficiency, and accuracy will increase—allowing the production of longer and more definitive sequences (10).

CHAPTER IV

RESULTS

Culture-based methods

Analysis of bacterial loads by direct plating on Tryptic Soy Agar (TSA) revealed that most of the foods tested had detectable bacterial loads (Table 1). Of the 9 foods that were plated and incubated under aerobic conditions, 6 showed a measurable bacterial load: pre-packaged spinach, whole milk, medium cheddar cheese, 80/20 ground beef, Atlantic salmon, and boneless, skinless chicken breast. Each of the foods that displayed aerobic bacterial load, with the exception of milk, also displayed anaerobic bacterial loads. The three foods that showed no aerobic bacterial load were banana, whole wheat bread, and fat-free vanilla yogurt.

Most food products showed an anaerobic load equal to, if not slightly greater than, the calculated aerobic load. The only exceptions to this observation were the milk and ground beef samples, both of which showed an anaerobic load significantly less than the aerobic load. Aerobic bacterial load of food samples ranged from $6.7x10^2$ CFU/ml in milk to $3.7x10^8$ CFU/g in cheddar cheese. Anaerobic bacterial loads ranged from $2.3x10^3$ in ground beef to $3.2x10^8$ in cheddar cheese.

TABLE 1. Aerobic and anaerobic bacterial load on food products.

Food Product	***Aerobic Load***	***Anaerobic Load***
Spinach	$2.1x10^6$ CFU/g	$5.4x10^5$ CFU/g
Whole Milk	$6.7x10^2$ CFU/ml	Below detection limit
Cheddar Cheese	$3.7x10^8$ CFU/g	$3.2x10^8$ CFU/g
80/20 Ground Beef	$1.6x10^6$ CFU/g	$2.3x10^3$ CFU/g
Atlantic Salmon	$1.4x10^6$ CFU/g	$1.4x10^7$ CFU/g
Chicken Breast	$4.2x10^5$ CFU/g	$2.8x10^6$ CFU/g

TABLE 2. Identification of aerobic bacterial isolates to the Genus level.

Food Product	***Identified Bacterial Genera***[α]
Spinach	*Enterobacter* spp.
Whole Milk	*Pseudomonas* spp., ***Bacillus* spp.**,
Cheddar Cheese	*Pseudomonas* spp., *Microbacterium* spp., *Kocuria* spp., *Clavibacter* spp.
80/20 Ground Beef	***Brochothrix* spp.**, *Yersinia* spp.
Atlantic Salmon	*Pseudomonas* spp., *Micrococcus* spp., *Acinetobacter* spp., *Arthrobacter* spp., *Shewanella* spp., ***Brochothrix* spp.**
Chicken Breast	*Pseudomonas* spp.

[α] The bacterial groups marked in bold are members of the Firmicutes division. No identified isolates were members of the Bacteroidetes division.

Thirty-seven isolates were analyzed by the fatty acid methyl ester (FAME) procedure. All 37 isolates submitted for analysis by the FAME procedure produced bacterial identifications; however, the Similarity Index (SIM) was only high enough for accurate genus identification in 28 of the isolates. The 9 isolates with SIMs lower than 2 were not included in the data set and were marked as "Not conclusive".

Of the 37 isolates sampled, 5 isolates were identified as members of the Firmicutes division (Table 2). These bacterial groups were *Brochothrix* spp. and *Bacillus* spp. The whole milk sample contained 2 isolates from the *Bacillus* genus. The salmon sample contained 1 isolate and the ground beef sample contained 2 isolates from the *Brochothrix* genus. None of the identified isolates was a member of the Bacteroidetes division of bacteria. A comprehensive list of all identified isolates is shown in Appendix B.

Pyrosequencing results

The pyrosequencing data was matched with sequences in the 16S database. Initially, the majority of sequences were matched to the genus level, although there were many species identified in all of the food samples. For subsequent analysis, however, all identified microbial organisms were sorted into groups of genera. This method of analysis was hypothesized to increase accuracy of the data analysis, since it is a

phylogenetic evaluation. A complete listing of all sequences from each food sample is provided in Appendix C.

Ground beef

The pyrosequencing results for the 80/20 ground beef provided 1390 "hits" of recognized bacterial 16S sequences. Of these bacterial sequences, 110 were not included in the final analysis due to E-values above 10^{-60} (Table 3). These 110 disqualified sequences were inconclusively identified as being members of the *Opitutus*, *Thermomonas*, *Prevotella*, *Bacteroides*, and *Cylindrospermum* genera. A phylogenetic breakdown of the metagenomic content of the ground beef reveals that it is composed of 10.5% Bacteroidetes, 72.3% Firmicutes, 15.4% Proteobacteria, and 1.8% Actinobacteria (Fig. 1). A phylogenetic analysis limited to only include Bacteroidetes and Firmicutes shows that the ground beef contains 12.7% Bacteroidetes and 87.3% Firmicutes.

TABLE 3. Results of pyrosequencing analysis of 80/20 ground beef.

Genus	Phylum	Sequences
Anaerobiospirillum	Proteobacteria	20
Bacteroides	Bacteroidetes	57
Brochothrix	Firmicutes	3
Buttiauxella	Proteobacteria	2
Chryseobacterium	Bacteroidetes	2
Clostridium	Firmicutes	8
Desulfovibrio	Proteobacteria	2
Enterococcus	Firmicutes	11
Escherichia	Proteobacteria	18
Lactobacillus	Firmicutes	31
Lactococcus	Firmicutes	848
Micrococcus	Actinobacteria	2
Niastella	Bacteroidetes	2
Nocardioides	Actinobacteria	18
Peptoniphilus	Firmicutes	10
Photobacterium	Proteobacteria	28
Prevotella	Bacteroidetes	74
Propionibacterium	Actinobacteria	2
Proteus	Proteobacteria	91
Pseudomonas	Proteobacteria	26
Ruminococcus	Firmicutes	3
Staphylococcous	Firmicutes	12
Succinivibrio	Proteobacteria	6
Sutterella	Proteobacteria	4
Total		**1280**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

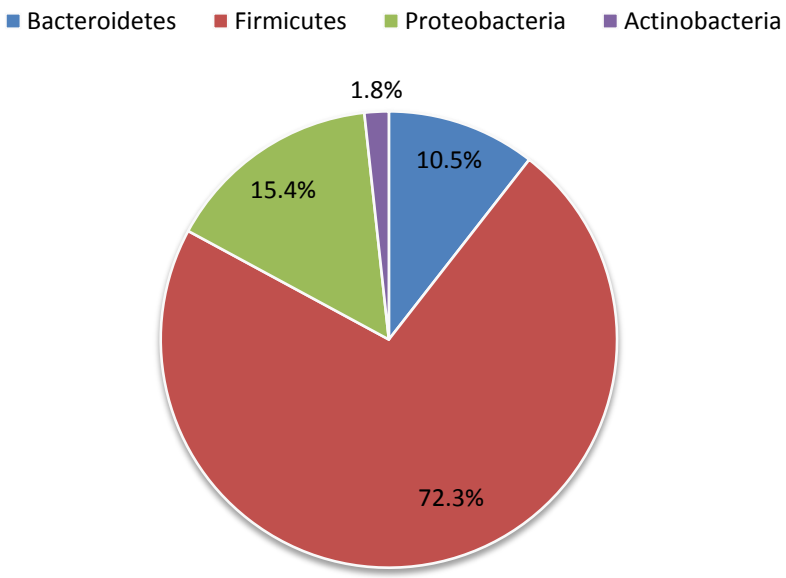

FIG. 1. Phylogenetic analysis of the metagenomic content of 80/20 ground beef determined by pyrosequencing.

Cheddar cheese

Pyrosequencing analysis of the metagenomic content of the medium cheddar cheese provided 4607 sequences identified with known bacterial 16S sequences (Table 4). The final analysis did not include 5 of these sequences due to insufficient E-values. The vast majority of microorganisms found in the cheese consisted of bacteria from the *Lactococcus* genus. (The disqualified sequences were also tentatively identified as members of the *Lactococcus* genus). Both the *Lactococcus* and *Streptococcus* genera are members of the Firmicutes phylum, suggesting that only Firmicutes were found in this food product.

Skinless chicken breast

The pyrosequencing analysis of the boneless, skinless chicken breast produced 1592 sequences, of which 282 sequences were rejected due to inconclusive E-values (Table

TABLE 4. Results of pyrosequencing analysis of medium cheddar cheese.

Genus	Phylum	Sequences
Lactococcus	Firmicutes	4600
Streptococcus	Firmicutes	2
Total		**4602**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

TABLE 5. Results of pyrosequencing analysis of boneless, skinless chicken breast.

Genus	Phylum	Sequences
Achromobacter	Proteobacteria	7
Acinetobacter	Proteobacteria	147
Aeromonas	Proteobacteria	22
Bacillus	Firmicutes	7
Bacteroides	Bacteroidetes	11
Buttiauxella	Proteobacteria	5
Castellaniella	Bacteroidetes	2
Enterococcus	Firmicutes	7
Escherichia	Proteobacteria	32
Flavobacterium	Bacteroidetes	8
Janthinobacterium	Proteobacteria	4
Micrococcus	Actinobacteria	3
Mitsuokella	Firmicutes	3
Nocardioides	Actinobacteria	157
Peptoniphilus	Firmicutes	3
Prevotella	Bacteroidetes	209
Proteus	Proteobacteria	96
Pseudomonas	Proteobacteria	465
Psychrobacter	Proteobacteria	52
Roseburia	Firmicutes	2
Shewanella	Proteobacteria	21
Succinivibrio	Proteobacteria	34
Sutterella	Proteobacteria	5
Variovorax	Proteobacteria	8
Total		**1310**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

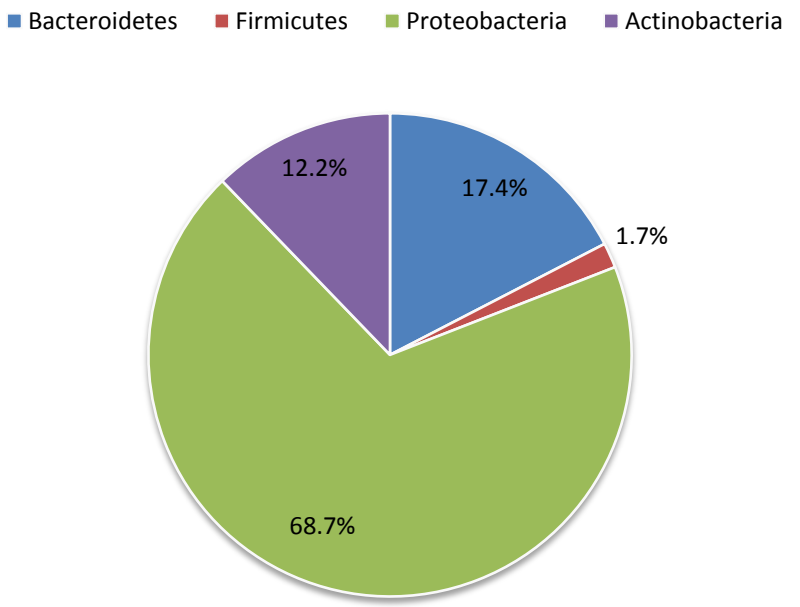

FIG. 2. Phylogenetic analysis of the metagenomic content of boneless, skinless chicken breast demonstrated by pyrosequencing.

5). The rejected "hits" were sequences from the *Prevotella*, *Cylindrospermum, Vibrio, Bacteroides*, *Coprothermobacter*, *Porphyromonas*, *Parabacteroides,* and *Ureibacillus* genera. The phylogenetic analysis of the chicken breast reveals a composition of 17.4% Bacteroidetes, 1.7% Firmicutes, 68.7% Proteobacteria, and 12.2% Actinobacteria (Fig. 2). If the scope of the phylogenetic breakdown is limited to compare only Bacteroidetes and Firmicutes, the microbial inhabitants of the chicken are 91.2% Bacteroidetes and 8.8% Firmicutes.

Atlantic salmon

The data produced by a pyrosequencing analysis of the Atlantic salmon was composed of 1261 identified sequences (Table 6). None of these sequences were excluded from the final analysis as a result of an insufficiently low E-value. On the contrary, the 16S sequences in the salmon sample were consistently very low, suggesting that all of the

TABLE 6. Results of pyrosequencing analysis of Atlantic salmon.

Genus	Phylum	Sequences
Acinetobacter	Proteobacteria	82
Aeromonas	Proteobacteria	114
Brochothrix	Firmicutes	34
Buttiauxella	Proteobacteria	3
Carnobacterium	Firmicutes	4
Chryseobacterium	Bacteroidetes	12
Elizabethkingia	Bacteroidetes	3
Epilithonimonas	Bacteroidetes	3
Flavobacterium	Bacteroidetes	11
Janinthobacterium	Proteobacteria	96
Lactococcus	Firmicutes	11
Photobacterium	Proteobacteria	28
Pseudomonas	Proteobacteria	670
Psychrobacter	Proteobacteria	31
Rahnella	Proteobacteria	44
Shewanella	Proteobacteria	107
Vibrio	Proteobacteria	10
Total		**1263**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

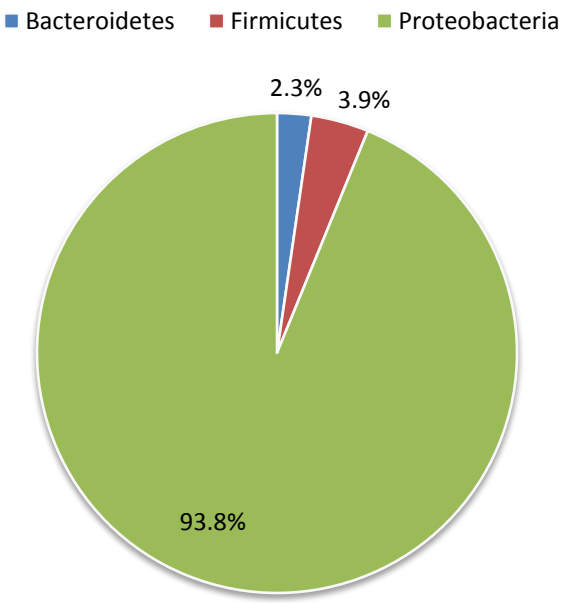

FIG. 3. Phylogenetic analysis of the metagenomic content of Atlantic salmon determined by pyrosequencing.

sequences found in the food sample were well-matched with the sequences of known organisms. A phylogenetic analysis of the metagenomic data reveals that the salmon contains 2.3% Bacteroidetes, 3.9% Firmicutes, and 93.8% Proteobacteria (Fig. 3). A comparison limited only to Bacteroidetes and Firmicutes results in a composition of 37.2% Bacteroidetes and 62.8% Firmicutes.

Whole milk

Pyrosequencing analysis of the whole milk produced a total of 3133 identified sequences, 2216 sequences of which were not included in the final analysis as a result of E-values above 10^{-80} (Table 7). The vast majority of excluded sequences were unidentified organisms with chloroplast-like 16S sequences (2097 sequences). The remaining excluded sequences were tentatively identified as members of the *Prevotella*, *Tropheryma*, *Bacteroides*, *Microlunatus*, *Hallella*, *Prosthecobacter*, *Cylindrospermum*,

TABLE 7. Results of pyrosequencing analysis of whole milk.

Genus	Phylum	Sequences
Anoxybacillus	Firmicutes	282
Bacillus	Firmicutes	9
Bacteroides	Bacteroidetes	16
Bifidobacterium	Actinobacteria	2
Clostridium	Firmicutes	6
Desulfovibrio	Proteobacteria	5
Escherichia	Proteobacteria	11
Friedmanniella	Actinobacteria	2
Fulvimonas	Proteobacteria	3
Fusobacterium	Fusobacterium	3
Lactococcus	Firmicutes	89
Pantoea	Proteobacterium	10
Prevotella	Bacteroidetes	110
Pseudomonas	Proteobacterium	340
Roseburia	Firmicutes	2
Ruminococcus	Firmicutes	5
Streptococcus	Firmicutes	2
Succinivibrio	Proteobacterium	18
Veillonella	Firmicutes	2
Total		**917**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

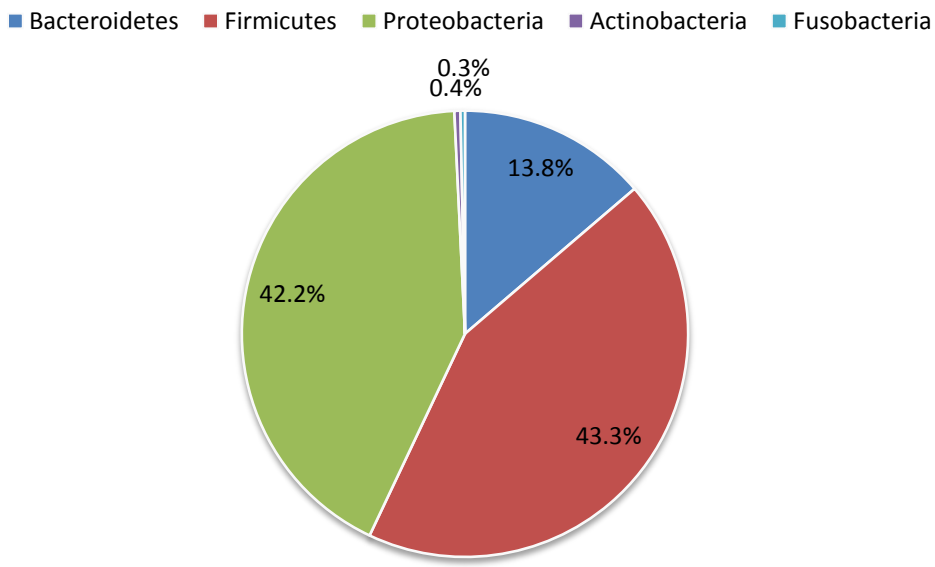

FIG. 4. Phylogenetic analysis of the metagenomic content of whole milk determined by pyrosequencing.

Clostridium, *Streptomonospora*, *Chondromyces*, and *Caloramator* genera. Phylogenetic analysis revealed that the metagenomic content of the whole milk was approximately 13.8% Bacteroidetes, 43.3% Firmicutes, 42.2% Proteobacteria, 0.4% Actinobacteria, and 0.3% Fusobacteria (Fig. 4). A phylogenetic analysis limited only to Bacteroidetes and Firmicutes showed a distribution of 24.1% Bacteroidetes and 75.9% Firmicutes.

Spinach

A pyrosquencing analysis of the spinach sample produced 2096 sequences that were identifiable to known microorganisms. From this data, 2046 sequences were rejected from the final data analysis due to insufficient E-values (Table 8). The vast majority of these excluded sequences (2018 sequences) were roughly identified as members of the *Cylindrospermum* genus, specifically the organism *Cylindrospermum stagnale*. The remainder of the excluded sequences were roughly correlated with organisms in the *Oscillatoria*, *Tropheryma*, and *Chlorogloeopsis* genera. All of the genera identified with excluded 16S sequences, except *Tropheryma*, were members of the phylum Cyanobacteria. All of the sequences subjected to final analysis were genera included in the phylum Proteobacteria. Therefore, no comparison of Bacteroidetes and Firmicutes could be performed.

TABLE 8. Results of pyrosequencing analysis of spinach.

Genus	Phylum	Sequences
Pantoea	Proteobacteria	16
Pseudomonas	Proteobacteria	32
Psychrobacter	Proteobacteria	2
Total		**50**

[α] Table only includes those genera identified by sequences with an E-value of 10^{-80} or lower.

Discussion

Culture-based identification

The plating results are significant because they illustrate that food samples harbor large numbers of aerobic and anaerobic microorganisms. An estimation of the anaerobic bacterial loads of food samples is critical because the vast majority of gut microbiota are either facultative or obligate anaerobes. It is interesting, then, that the majority of food products analyzed contain nearly equal amounts of aerobic and anaerobic bacteria. The only two foods that showed a significant difference were whole milk and ground beef. This difference could be due to a bias arising from the culture media that was used in these studies. Many bacterial organisms cannot proliferate on media that does not supply specific conditions and nutrients. So, even if the organism was viable (ie., culturable), it might not have been able to grow on the media used. This study employed only TSA for both the aerobic and anaerobic estimations. The numbers could have been vastly different if other media such as R2A, Plate Count Agar, or Brain-Heart Infusions Agar were used. The plating of multiple bacterial species on a single media may also have resulted in competition between bacterial colonies that skewed the results in favor of those species that grow well on the media. This difference highlights a major limitation of culture-based methods in that the results are highly dependent on the media that was used. Also, anaerobic bags were used for the enumeration of anaerobes. The use of a more specialized anaerobic hood could have resulted in different results. More importantly, however, the processing of foods could have resulted in the specific results

that were obtained. The pasteurization of milk may have eliminated temperature-sensitive anaerobic organisms. It should be noted that the anaerobes were not characterized as being either obligate or facultative anaerobes. The use of an anaerobic hood during the enumeration and subsequent FAME characterization could have provided some insight into their oxygen sensitivity. Overall, the results from the conventional microbiological analyses reveal that foods contain a large and diverse microbial load.

Bacteroidetes and Firmicutes

The overall objective of this research was to determine whether Bacteroidetes and Firmicutes, which had been linked to obesity in humans, were present in selected food samples. The results from this study provide strong evidence that our underlying hypothesis was correct. Though culture-based identification of bacterial isolates confirmed that 5 out of the 27 isolates were Firmicutes, the strongest supporting evidence was obtained by the extremely powerful pyrosequencing methodology. The different food samples harbored a significant diversity of organisms belonging to either Bacteroidetes or Firmicutes. The only exception was spinach, which did not contain any Bacteroidetes or Firmicutes. Both pasteurized milk and cheddar cheese samples were dominated by Firmicutes, although the milk sample also contained large numbers of bacterial groups from other phyla. There was no clear correlation between the phylogenetic proportions of the inhabitants of ground beef, chicken breast, and salmon. The chicken breast was the only sample that had a majority of Bacteroidetes as

compared to Firmicutes. Several of the food products contained large populations of other phyla such as Proteobacteria and Actinobacteria. However, more than 99% of gut bacteria are either Bacteroidetes or Firmicutes (6).

Human gut flora identified in pyrosequencing data

This experiment not only found Bacteroidetes and Firmicutes in the food samples, but also found specific bacterial groups within these phyla that are known inhabitants of the human gut (Table 9). The gut inhabitants include organisms from *Bacteroides spp.*, *Bifidobacterium spp.*, *Eubacterium spp.*, *Clostridium spp.*, *Ruminococcus spp.*, *Escherichia spp.*, *Enterobacter spp.*, *Lactobacillus spp.*, and *Proteus spp*. This finding is noteworthy because it implies that consumption of a particular type of food could affect the microbial diversity of the gastrointestinal tract. More directed studies, however, are needed to validate this claim.

The data was also analyzed to determine if the phylogenetic proportions of each of the food products would change drastically if limited only to organisms that have be proven to inhabit the human intestinal tract (Fig. 5). The ground beef contained 1147 sequences that were similar to human gut microbiota. Of these, 11.4% were Bacteroidetes, 78.6% were Firmicutes, 9.8% were Proteobacteria, and 0.2% were Actinobacteria. The cheddar cheese sample contained 4602 sequences, all of which were members of the Firmicutes phylum. The chicken breast sample contained 367 sequences, consisting of 59.9% Bacteroidetes, 3.8% Firmicutes, and 36.2% Proteobacteria. The Atlantic salmon

TABLE 9. Human gut microbiota found in pyrosequencing analysis of food samples.

Food Product	Bacteroidetes	Firmicutes	Proteobacteria	Actinobacteria
Beef	*Bacteroides spp., Prevotella spp.*	*Clostridium spp., Enterococcus spp., Lactobacillus spp., Lactococcus spp., Ruminococcus spp.*	*Escherichia spp., Proteus spp., Sutterella spp.*	*Propionibacterium spp.*
Cheese	none	*Lactococcus spp., Streptococcus spp.*	none	none
Chicken	*Bacteroides spp., Prevotella spp.*	*Bacillus spp., Enterococcus spp.*	*Escherichia spp., Proteus spp., Sutterella spp.*	none
Fish	none	*Lactococcus spp.*	none	none
Milk	*Bacteroides spp., Prevotella spp.*	*Bacillus spp., Clostridium spp., Lactococcus spp., Ruminococcus spp., Streptococcus spp.*	*Escherichia spp.*	*Bifidobacterium spp.*
Spinach	none	none	none	none

[α] Organisms are sorted by phyla. All bacterial groups identified to genus level.

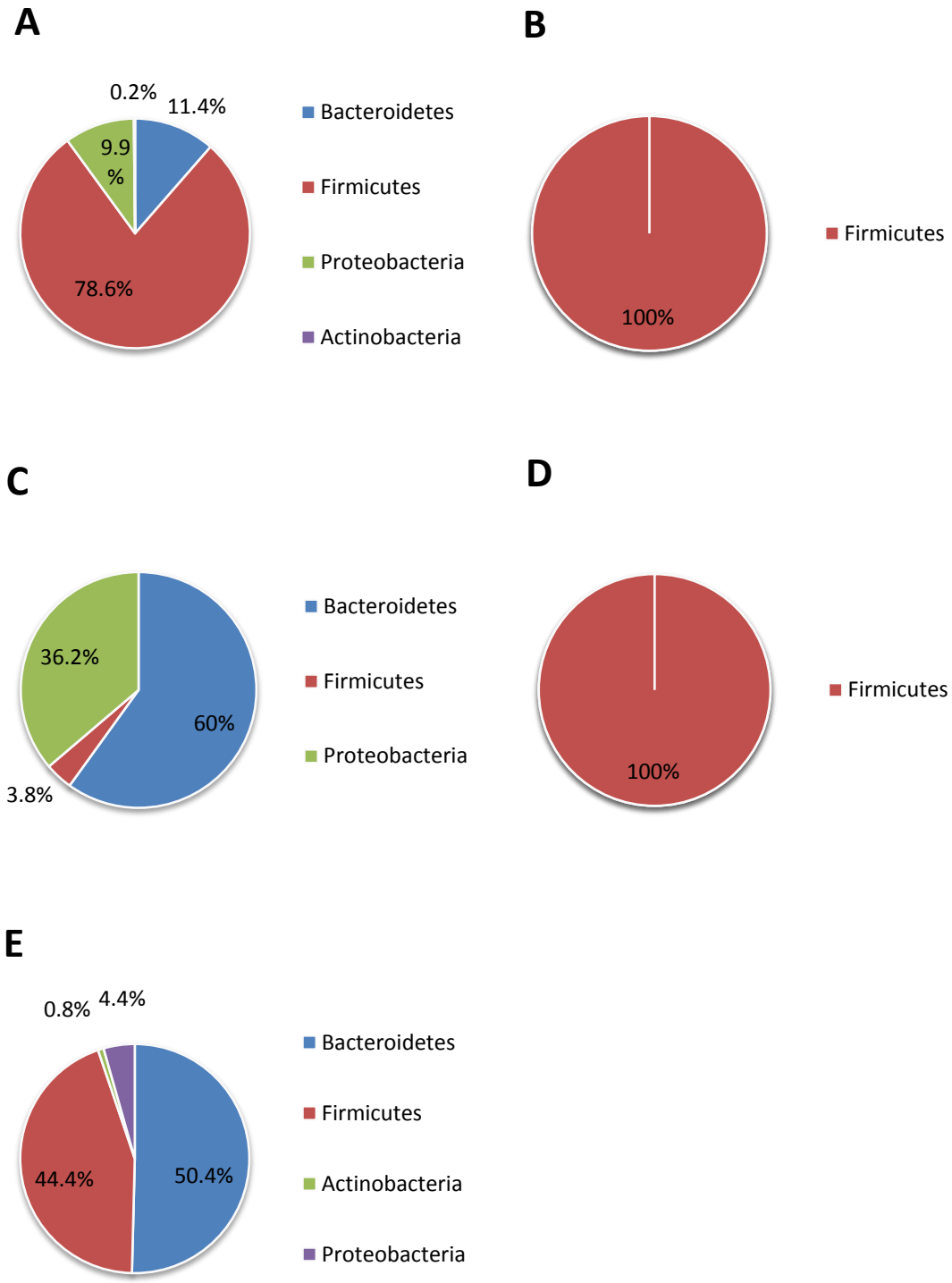

FIG. 5. Proportions of known gut microbiota present in food products. (A) 80/20 Ground Beef (B) Medium Cheddar Cheese (C) Boneless, Skinless Chicken Breast (D) Atlantic Salmon (E) Whole Milk

contained 11 sequences consistent with human gut microbiota, all of which were members of the Firmicutes phylum. The whole milk contained 237 sequences, consisting of 50.4% Bacteroidetes, 44.4% Firmicutes, 4.4% Proteobacteria, and 0.8% Actinobacteria. The spinach contained none of the microorganisms that were identified to be similar to organisms found in the human gut.

When limited to only those bacterial groups found in the human gut, the chicken breast sample showed an increase in the proportion of Bacteroidetes (17.4% to 59.9%) and Firmicutes (1.7% to 3.8%) and a decrease in the proportion of Proteobacteria (68.7% to 36.2%). Both of the genera found in the cheddar cheese were also found in the human gut. The ground beef sample showed a slight increase in the proportion of Bacteroidetes (10.5% to 11.4%) and Firmicutes (72.3% to 78.6%) and a decrease in the proportion of Proteobacteria (15.4% to 9.8%) and Actinobacteria (1.8% to 0.2%). The only bacterial groups found in the Atlantic salmon that were previously shown to be gut inhabitants were the Firmicutes. Milk showed a significant increase of Bacteroidetes (13.8% to 50.4%) when limited to only the gut microbiota. The proportion of Firmicutes increased slightly (42.2% to 44.4%) and the proportion of Proteobacteria decreased dramatically (42.4% to 4.4%).

The finding that these gut microbiota are present in several of the examined food products does not prove that they help to colonize or otherwise influence the bacteria of the gut. Many factors, including survival of these bacterial populations during passage

through the highly acidic stomach environment, consumption patterns of particular type of foods, food turnover in the gastrointestinal tract, and the physiological conditions of the host will probably influence the microbial diversity of the gastrointestinal tract.

Correlation between nutritional content and bacterial load

After compilation of the pyrosequencing data, the total percentage of Bacteroidetes and Firmicutes in each food was compared to select elements of that food's nutritional information including calories per gram of food product, total fat per gram of food product, and total carbohydrates per gram of food product. The purpose of this analysis was to determine if there was a correlation between the percent of either Bacteroidetes or Firmicutes in a food product and the nutritional content of that food. The milk data was not included in the correlation calculation because nutritional information could not be compared in terms of grams. A regression analysis of the data showed no correlation between percentage of Bacteroidetes and either calories (R^2=0.0171), total fat (R^2=0.0680), or total carbohydrates (R^2=0.1933) in the food products. However, a regression analysis comparing nutritional content to percentage of Firmicutes in the food products did show some evidence of a correlation. A comparison of overall percentage of Firmicutes in each food with the food's calories per gram produced an R^2 value of 0.8047 (Fig. 6). A comparison of overall percentage of Firmicutes in each food with the food's fat grams per gram produced an R^2 value of 0.9575 (Fig. 7). No correlation was observed between overall percentage of Firmicutes and total carbohydrates (R^2=0.1611) in each food.

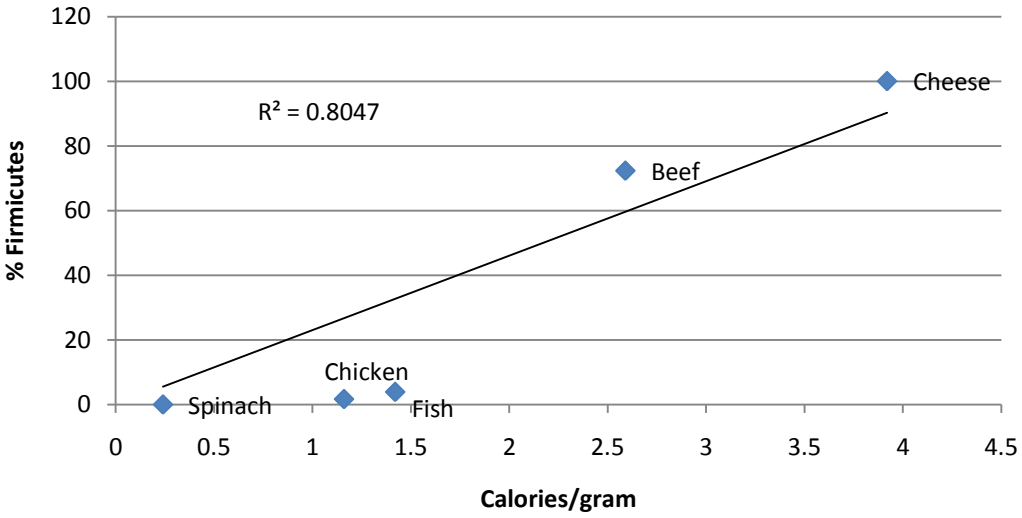

FIG. 6. Correlation between calories contained per gram and overall percentage of Firmicutes in each food product.

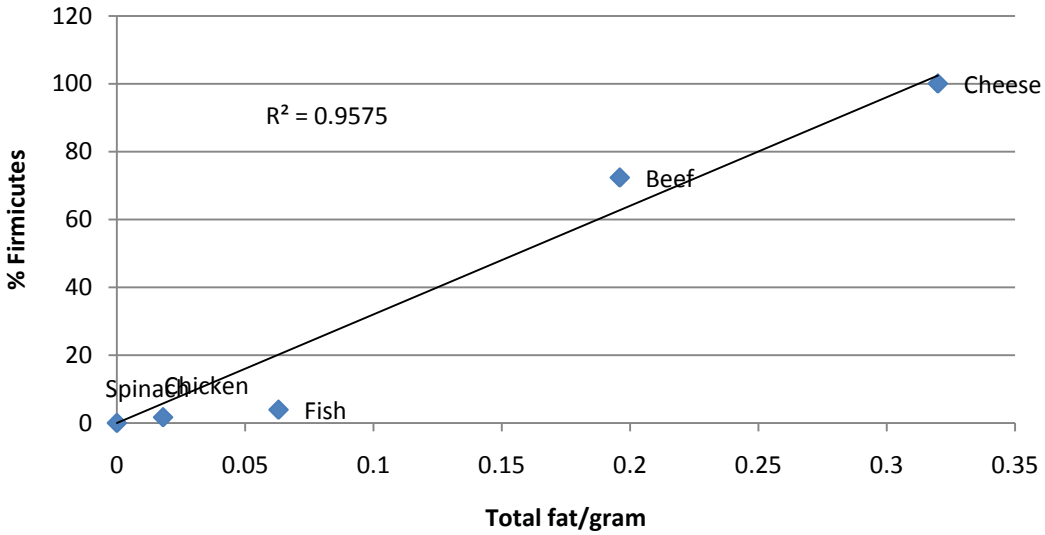

FIG. 7. Correlation between total fat per gram and overall percentage of Firmicutes in each food product.

The data collected by the molecular analysis provides slight evidence to suggest that there may be a correlation between overall percentage of Firmicutes and nutritional content. This assertion, that there is a correlation between the overall percentage of Firmicutes and the number of calories or fat grams contained in a food, will, however, require much further experimentation and analysis. If there is, in fact, a relationship between these two variables, it could explain why a higher proportion of Firmicutes are present in obese individuals. On a rather simplistic level, obese individuals are more likely to consume a diet high in calories or fat. This consumption thereby could increase the percentage of Firmicutes entering the gastrointestinal tract. It is also possible that more fat or calories in the diet could simply create a more favorable intestinal environment for Firmicutes to proliferate.

Comparison of culture-based methods vs. molecular methods

An evaluation of the data obtained by the culture-based and molecular approaches shows that these two methods display a staggering disparity in the results obtained by each method. The culture-based methods involved the analysis of 37 isolates, of which 29 were identified. The isolates identified by the FAME analysis in ground beef were members of *Brochothrix spp.* and *Yersinia spp.* Sequences consistent with the *Brochothrix* genus were also found in the group beef pyrosequencing results, but sequences identified with the *Yersinia* genus were not. The cultured isolates identified in cheddar cheese were *Pseudomonas spp.* and *Kocuria spp.* Neither of these organisms, however, was found in the pyrosequencing data for cheese. The isolates identified in

chicken breast were members of *Pseudomonas spp.* Sequences belonging to this genus were identified in the chicken breast pyrosequencing data. *Pseudomonas spp.*, *Actinetobacter spp.*, *Kocuria spp.*, *Shewanella spp.*, and *Brochothrix spp,* and *Micrococcus spp.* were all identified isolates in the culture-based analysis of the Atlantic salmon. All of these genera were identified in the salmon pyrosequencing data except for *Kocuria spp*. and *Micrococcus spp*. The isolates taken from the whole milk were identified as members of the *Bacillus* and *Pseudomonas* genera and both of these genera were also detectable by the pyrosequencing method. *Pantoea spp*. were detectable in spinach by both the FAME and pyrosequencing methods. Out of 28 isolates detected in the food samples by culture-based methods, 19 were identified in the corresponding pyrosequencing data. In contrast, the pyrosequencing method provided 14,081 bacterial 16S sequences, of which 9,422 sequences were included in the final analysis. Only a percentage of the bacterial genera on the food samples were identifiable by the culture-based methods. Approximately 18.3% of the pyrosequencing data was present in the culture-based results.

The aerobic and anaerobic culturing of the food products, isolation, and identification procedures took several weeks. Once the DNA extractions were performed, the pyrosequencing analysis was completed in a few days. Pyrosequencing was, however, costly. Also, the pyrosequencing results do not provide information regarding whether the 16S sequences originated from live or dead cells in the food sample. Because of the nature of the FAME identification procedure, only aerobic organisms were able to be

analyzed. This prevented any anaerobes from being identified by culture-based methods. It is unknown why this discrepancy exists between the culture-based and the molecular results. It is possible that they may be explained by limitations in the database used for the FAME analysis. It is also possible that the pyrosequencing run did not analyze all of the 16S sequences present in the loaded sample.

Source of identified microorganisms

The types of organisms found in a food product could depend on several variables including: the natural flora found on a specific plant or animal, the plant or animal's health, and the processing and packaging conditions encountered by the food sample. It would be a challenge to determine whether the organisms found in this study originated pre-harvest or during post-harvest processing and handling. Directed feeding studies with genetically well-characterized microbial strains could shed some light on this important question. It must be pointed out, however, that irrespective of the source of these organisms, the consuming public are being exposed to these populations.

Medical significance

Several bacterial groups of medical significance were found in the molecular analysis of the food-associated microbial populations. Pyrosequencing has been used previously for the identification of pathogenic bacteria (12). Most of the analyzed food samples contained genera that have been known to cause disease or opportunistic infections. These include *Vibrio spp.*, *Proteus spp*. *Aeromonas spp.*, *Clostridium spp.*, *Pseudomonas*

spp., and *Acinetobacter spp.* (18). *Aeromonas hydrophila*, identified in the poultry and fish samples, has been identified as a major human pathogen capable of causing gastroenteritis and wound infections (11). Anaerobic organisms cause between 1% and 3% of all bacteremic infections (18). Organisms from the *Bacteroides* and *Prevotella* genera have been known to cause sinus and periodontal infections, intraaabdominal infections, and gynecologic infections—mostly in immunocompromised persons or as a result of trauma (18).

Unidentified bacterial 16S sequences

As was reported in the results, a portion of the bacterial 16S sequences reported in the pyrosequencing data were not utilized in this research because of their lower correlation to known organisms, represented by a relatively high E-value. The majority of these bacterial sequences were tentatively identified as members of the Bacteroidetes, mostly from the genera *Bacteroides* and *Prevotella*. The majority of organisms in the Bacteroidetes division are strict anaerobes and difficult to culture under the best of conditions. It is highly probable that the Bacteroidetes phylum contains many species and subspecies that have yet to be properly identified and characterized. This explanation is most likely also true for other higher E-value, non-Bacteroidetes sequences found in this study. As newly discovered organisms are defined and their 16S sequences added to genomic databases, these organisms will be able to be positively identified.

CHAPTER V

SUMMARY

The major findings of this study are as follows:

1. **Microbiological analysis of the food samples on general purpose media (TSA) showed significant aerobic and anaerobic bacterial loads.** The bacterial load in the solid foods ranged from $4.2x10^5$ CFU/g to $3.7x10^8$ CFU/g for aerobic bacteria and from $2.3x10^3$ CFU/g to $3.2x10^8$ CFU/g for anaerobic bacteria. The milk samples showed no anaerobic load, but an aerobic load of $6.7x10^2$ CFU/ml.
2. **Aerobic isolates taken from each of the food cultures and identified by FAME analysis led to the identification of 27 isolates to the genera level.** Of these, 5 belonged to the Firmicutes phylum. None of the bacterial isolates were identified as belonging to the Bacteroidetes phylum.
3. **Metagenomic analysis of the food products by pyrosequencing revealed significant populations of organisms belonging to both the Bacteroidetes and Firmicutes phylums, as well as organisms from other bacteria phyla.** All microbial groups included in the final analysis belonged to one of the following phyla: Actinobacteria, Bacteroidetes, Firmicutes, Fusobacteria, or Proteobacteria. The ground beef, Atlantic salmon, skinless chicken breast, and milk samples showed an extremely diverse microbial population.
4. **Each food type appears to have its own unique bacterial population**. The chicken breast was dominated by organisms from the phylum Proteobacteria, with lesser

numbers of Acinobacteria, and Bacteroidetes. Spinach was inhabited by organisms included in the Proteobacteria phylum. Only two genera were found in the cheddar cheese—both Firmicutes. The ground beef sample contained mostly Firmicutes, with lesser numbers of Proteobacteria, Bacteroidetes, and Actinobacteria. The whole milk sample contained nearly equal numbers of Proteobacteria and Firmicutes, with lesser numbers of Bacteroidetes and Actinobacteria. Proteobacteria dominated the salmon, with very few organisms from the Bacteroidetes and Firmicutes phyla.

5. **Correlation of the pyrosequencing results with known gut flora also found specific organisms in the food samples that are also present in the human intestinal tract.** All of the analyzed food products except spinach contained genera that have been identified as human gut microbiota.
6. **Comparison of the results obtained by culture-based methods and pyrosequencing showed some discrepancies.** Of the 27 colony isolates identified by the FAME method, 19 were identified in the correlating pyrosequencing data. Approximately 18.3% of the pyrosequencing data was also present in the culture-based results.
7. **Analysis of the relationship between metagenomic content and nutritional content revealed slight evidence of a correlation between overall percentage of Firmicutes in a food product and that food's fat and/or calories per gram.**

Future research directions

The results of these experiments have shed light on a number of interesting possibilities. Further research is needed to gain a better understanding of the relationship between the microbial populations on foods and human obesity. Specifically, foods need to be categorized into different classes based on the demographics of those consuming such foods (infant foods, typical pre-teen to young adult foods, etc.) and the microbial diversity of such foods be delineated. Additionally, detailed studies are needed to track the changes in the microbial diversity of the gastrointestinal tract from infancy to adulthood and attempt to correlate it to the diversity in the foods consumed during the same periods. Extraction and analysis of metagenomic data should be repeated for a wider variety of food products to determine if microbial contents are widely different or whether there is a pattern in the types of microorganisms that may be present. Similarly, a better understanding of the changes in microbial diversity from the time the food is packaged to the time of consumption is important. Since the pyrosequencing technique is a relatively new technology, the capabilities and efficiency of the pyrosequencing and 16S cloning-based protocols for metagenomic analysis should be compared to determine strengths and weakness associated with each technique. Improved techniques for the isolation and culture of anaerobic bacteria are needed, especially those bacteria identified as a part of the anaerobic load of food products. Research in the last few years has significantly improved our knowledge of human gut microbes, but more is still needed. The proven variability of gut microbiota between individuals makes it likely that there are still species and genera that have yet to be discovered. Lastly, it is apparent

from this research that many different organisms inhabit the evaluated food products. It is unclear, however, the sources of these bacteria on these foods. Further research is needed to evaluate which organisms are present in the original plant or animal and which were deposited during processing and packaging.

REFERENCES

1. **Ahmadian, A., M. Ehn, and S. Hober**. 2006. Pyrosequencing: history, biochemistry and future. Clinica Chimica Acta. **363**:83-94.

2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman**. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403-410.

3. **Backhed, F., H. Ding, T. Wang, L. V. Hooper, G. Y. Koh, A. Nagy, C. F. Semenkovich, and J. I. Gordon.** 2004. The gut microbiota as an environmental factor that regulates fat storage. Proc. Natl. Acad. Sci. U.S.A. **101**:15718-15723.

4. **Bajzer, M., and R. J. Seeley.** 2006. Obesity and gut flora. Nature. **444**:1009-1010.

5. **Eckburg, P. B., E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, and D. A. Relman.** 2005. Diversity of the human intestinal microbial flora. Science. **308**:1635-1638.

6. **Gill, S. R., M. Pop, R. T. DeBoy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, and K. E. Nelson.** 2006. Metagenomic analysis of the human distal gut microbiome. Science. **312**:1355-1359.

7. **Guarner, F. and J. R. Malagelada**. 2003. Gut flora in health and disease. The Lancet. **360**:512-519.

8. **Hayashi, H., M. Sakamoto, and Y. Benno.** 2002. Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic and culture-based methods. Microbiol. Immunol. **46**:535-548.

9. **Huang, X., and A. Madan**. 1999. CAP3: A DNA sequencing assembly program. Genome Res. **9**:868-877.

10. **Huse, S. M., J. A. Huber, H. G. Morrison, M. L. Sogin, and D. M. Welch**. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biology. **8**:R143.

11. **Janda, J. M. and S. L. Abbott**. 1998. Evolving concepts regarding the genus Aeromonas: an expanding panorama of species, disease presentations, and unanswered questions. Clinical Infectious Diseases. **27**:332-344.

12. **Jonasson, J., Olofsson, M., and H. Monstein**. 2002. Classification, identification, and subtyping of bacteria based on pyrosequencing and signature matching of 16S rDNA fragments. APMIS. **110**:263-272.

13. **Kopelman, P. G.** 2000. Obesity as a medical problem. Nature**. 404**:635-643.

14. **Ley, R. E., F. Backhed, P. Turnbaugh, C. A. Lozupone, R. D. Knight, and J. I. Gordon.** 2005. Obesity alters gut microbial ecology. Proc. Natl. Acad. Sci. U.S.A. **102**:11070-11075.

15. **Ley, R. E., P. J. Turnbaugh, S. Klein, and J. I. Gordon.** 2006. Human gut microbes associated with obesity. Nature. **444**:1022-1023.

16. **Liu, Z., C. Lozupone, M. Hamady, F. D. Bushman, and R. Knight**. 2007. Short pyrosequencing reads suffice for accurate microbial community analysis. Nucleic Acids Res. **35**:e120

17. **Maidak, B. L., J. R. Cole, T. G. Lilburn, C. T. Parker, P. R. Saxman, R. J. Farris, G. M. Garrity, G. J. Olsen, T. M. Schmidt, and J. M. Tiedje**. 2001. The RDP-II (Ribosomal database project). Nucleic Acids Res. **29**:173-174.

18. **Murray, P. R., R. S. Rosenthal, and M. A. Pfaller**. 2005. Medical microbiology, fifth edition. Elsevier, Philadelphia, PA.

19. **Orrhage, K., and C. E. Nord.** 1999. Factors controlling the bacterial colonization of the intestine in breastfed infants. Acta Pediatric Suppl. **430**:47-57.

20. **Roesch, L. F. W., R. R. Fulthorpe, A. Riva, G. Casella, A. K. M. Hadwin, A. D. Kent, S. H. Daroub, F. A. O. Camargo, W. G. Farmerie, and E. W. Triplett**. 2007. Pyrosequencing enumerates and contrasts soil microbial diversity. ISME. **1**:283-290.

21. **Ronaghi, M**. 2001. Pyrosequencing sheds light on DNA sequencing. Genome Research. **11**:3-11.

22. **Ronaghi, M., S. Karamohamed, B. Pettersson, M. Uhlen, and P. Nyren**. 1996. Real-time DNA sequencing using detection of pyrophosphate release. Anal. Biochem. **242**:84-89.

23. **Sears, C. L.** 2005. A dynamic partnership: celebrating our gut flora. Anaerobe. **11**:247-251.

24. **Suau, A., R. Bonnet, M. Sutren, J. Godon, G. R. Gibson, M. D. Collins, and J. Doré.** 1999. Direct analysis of genes encoding 16S rRNA from complex

communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. **65**:4799-4807.

25. **Tuohy, M. J., G. S. Hall, M. Sholtis, and G. W. Procop**. 2005. Pyrosequencing™ as a tool for the identification of common isolates of *Mycobacterium* sp. Diagn. Microbiol. Infect. Dis. **51**:245-250.

26. **Turnbaugh, P. J., R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, and J. I. Gordon.** 2006. An obesity-associated gut microbiome with increased capacity for energy harvest. Nature. **444**:1027-1031.

APPENDIX A

FOOD PRODUCT NUTRITIONAL INFORMATION

TABLE 10. Nutritional information for analyzed food samples.

Yoplait Original 99% Fat-Free French Vanilla Flavored Yogurt (170 g)		
Serving Size: 1 Container	Calories: 170	Calories from Fat: 15
Total Fat: 1.5 g	Saturated Fat: 1 g	Fiber: 0 g
Total Carbohydrate: 33 g	Sugars: 27 g	Protein: 5 g
H-E-B Medium Cheddar Cracker Cut Cheese (170 g)		
Serving Size: 4 pieces (28 g)	Calories: 110	Calories from Fat: 80
Total Fat: 9 g	Saturated Fat: 5 g	Fiber: 0 g
Total Carbohydrate: <1 g	Sugars: 0 g	Protein: 7 g
H-E-B 80/20 Ground Beef (1.09 lb)		
Serving Size: 4 oz. (112 g)	Calories: 290	Calories from Fat: 200
Total Fat: 22 g	Saturated Fat: 9 g	Fiber: 0 g
Total Carbohydrate: 0 g	Sugars: 0 g	Protein: 21 g
Hill Country Fare Wheat Enriched Bread (680 g)		
Serving Size: 1 slice (28 g)	Calories: 80	Calories from Fat: 5
Total Fat: 0.5 g	Saturated Fat: 0 g	Fiber: 2 g
Total Carbohydrate: 15 g	Sugars: 2 g	Protein: 3 g
Fresh Express Spinach (255 g)		
Serving Size: 85 g	Calories: 20	Calories from Fat: 0
Total Fat: 0 g	Saturated Fat: 0 g	Fiber: 2 g
Total Carbohydrate: 3 g	Sugars: 0 g	Protein: 2 g

TABLE 10 Continued

Sanderson Farms 97% Fat-Free Boneless, Skinless Breast Fillets (1.39 lbs)		
Serving Size: 4 oz. (112 g)	Calories: 130	Calories from Fat: 20
Total Fat: 2 g	Saturated Fat: 0.5 g	Fiber: 0 g
Total Carbohydrate: 0 g	Sugars: 0 g	Protein: 26 g
Oak Farms Vitamin D Milk (64 fl oz)		
Serving Size: 1 cup	Calories: 150	Calories from Fat: 70
Total Fat: 8 g	Saturated Fat: 5 g	Trans Fat: 0 g
Total Carbohydrate: 11 g	Sugars: 11 g	Protein: 8 g
H-E-B "Organic" Banana		
Serving Size: 1 medium banana (118g)	Calories: 105	Calories from Fat: Not Available
Total Fat: 0.4 g	Saturated Fat: Not Available	Trans Fat: Not Available
Total Carbohydrate: 27 g	Sugars: 14.4 g	Protein: 1.3 g
H-E-B Fresh Atlantic Salmon Fillet (0.65 lb)		
Serving Size: 1 package	Calories: 417	Calories from Fat: Not Available
Total Fat: 18.6 g	Saturated Fat: Not Available	Trans Fat: Not Available
Total Carbohydrate: 0 g	Sugars: 0 g	Protein: 58.3 g

[α] Nutritional information for the yogurt, cheddar cheese, ground beef, wheat bread, spinach, chicken breast, and whole milk was taken from product packaging. Nutritional information for the banana and Atlantic salmon was not listed on the packaging and was therefore taken from the USDA Nutrient Data Laboratory at http://www.nal.usda.gov/fnic/foodcomp/search/.

APPENDIX B

IDENTIFICATION OF AEROBIC ISOLATES

TABLE 11. Identity of aerobic isolates as determined by the FAME procedure.

Isolate Number	Food Sample (Dilution)	Bacterial Genus	Bacterial Division
1	Chicken (10^{-3})	*Pseudomonas*	Proteobacteria
2	Chicken (10^{-3})	*Pseudomonas*	Proteobacteria
3	Chicken (10^{-3})	*Pseudomonas*	Proteobacteria
4	Chicken (10^{-3})	*Pseudomonas*	Proteobacteria
5	Chicken (10^{-3})	*Pseudomonas*	Proteobacteria
6	Spinach (10^{-3})	*Enterobacter*	Proteobacteria
7	Spinach (10^{-3})	*Enterobacter*	Proteobacteria
8	Cheese (10^{-1})	*Pseudomonas*	Proteobacteria
9	Cheese (10^{-1})	*Microbacterium*	Actinobacteria
10	Cheese (10^{-1})	*Pseudomonas*	Proteobacteria
11	Cheese (10^{-2})	*Kocuria*	Actinobacteria
12	Cheese (10^{-3})	*Clavibacter*	Actinobacteria
13	Cheese (10^{-3})	*Pseudomonas*	Proteobacteria
14	Beef (10^{-4})	*Not conclusive*	n/a
15	Beef (10^{-4})	*Not conclusive*	n/a
16	Beef (10^{-3})	*Brochothrix*	Firmicutes
17	Beef (10^{-3})	*Brochothrix*	Firmicutes
18	Beef (10^{-3})	*Not conclusive*	n/a
19	Beef (10^{-2})	*Not conclusive*	n/a
20	Beef (10^{-2})	*Not conclusive*	n/a
21	Beef (10^{-2})	*Yersinia*	Proteobacteria
22	Fish (10^{-4})	*Micrococcus*	Actinobacteria
23	Fish (10^{-4})	*Not conclusive*	n/a
24	Fish (10^{-4})	*Pseudomonas*	Proteobacteria
25	Fish (10^{-4})	*Acinetobacter*	Proteobacteria

Table 11 Continued

26	Fish (10^{-4})	*Not conclusive*	n/a
27	Fish (10^{-4})	*Pseudomonas*	Proteobacteria
28	Fish (10^{-3})	*Arthrobacter*	Actinobacteria
29	Fish (10^{-3})	*Shewanella*	Proteobacteria
30	Fish (10^{-3})	*Brochothrix*	Firmicutes
34	Milk (10^{-1})	*Not conclusive*	n/a
36	Milk (10^{-1})	*Bacillus*	Firmicutes
37	Milk (10^{-1})	*Not conclusive*	n/a
38	Milk (10^{-1})	*Pseudomonas*	Proteobacteria
39	Milk (10^{-1})	*Bacillus*	Firmicutes
40	Milk (10^{-1})	*Pseudomonas*	Proteobacteria

APPENDIX C

COMPLETE PYROSEQUENCING RESULTS FOR EACH FOOD SAMPLE

TABLE 12. Complete pyrosequencing results for analyzed food samples.

Ground Beef			
Genus	**Sequences**	**Genus**	**Sequences**
Anaerobiospirillum spp.	20	*Nocardioides spp.*	18
Bacteroides spp.	98	*Opitutus spp.*	7
Brochothrix spp.	3	*Peptoniphilus spp.*	10
Buttiauxella spp.	2	*Photobacterium spp.*	28
Chryseobacterium spp.	2	*Prevotella spp.*	112
Clostridium spp.	8	*Propionibacterium spp.*	2
Cylindrospermum spp.	20	*Proteus spp.*	91
Desulfovibrio spp.	2	*Pseudomonas spp.*	26
Enterococcus spp.	11	*Ruminococcus spp.*	3
Escherichia spp.	18	*Staphylococcous spp.*	12
Lactobacillus spp.	31	*Succinivibrio spp.*	6
Lactococcus spp.	848	*Sutterella spp.*	4
Micrococcus spp.	2	*Thermomonas spp.*	4
Niastella spp.	2	**Total**	**1390**
Atlantic Salmon			
Genus	**Sequences**	**Genus**	**Sequences**
Acinetobacter spp.	82	*Janinthobacterium spp.*	96
Aeromonas spp.	114	*Lactococcus spp.*	11
Brochothrix spp.	34	*Photobacterium spp.*	28
Buttiauxella spp.	3	*Pseudomonas spp.*	670
Carnobacterium spp.	4	*Psychrobacter spp.*	29
Chryseobacterium spp.	12	*Rahnella spp.*	44
Elizabethkingia spp.	3	*Shewanella spp.*	107
Epilithonimonas spp.	3	*Vibrio spp.*	10
Flavobacterium spp.	11	**Total**	**1261**

Table 12 Continued

Boneless, Skinless Chicken Breast			
Genus	**Sequences**	**Genus**	**Sequences**
Achromobacter spp.	7	*Nocardioides spp.*	157
Acinetobacter spp.	147	*Parabacteroides spp.*	20
Aeromonas spp.	22	*Peptoniphilus spp.*	3
Bacillus spp.	7	*Porphyromonas spp.*	12
Bacteroides spp.	96	*Prevotella spp.*	270
Buttiauxella spp.	5	*Proteus spp.*	96
Castellaniella spp.	2	*Pseudomonas spp.*	465
Coprothermobacter spp.	2	*Psychrobacter spp.*	52
Cylindrospermum spp.	88	*Roseburia spp.*	2
Desulfatibacillum spp.	4	*Shewanella spp.*	21
Enterococcus spp.	7	*Succinivibrio spp.*	34
Escherichia spp.	32	*Sutterella spp.*	5
Flavobacterium spp.	8	*Ureibacillus spp.*	7
Janthinobacterium spp.	4	*Variovorax spp.*	8
Micrococcus spp.	3	*Vibrio spp.*	3
Mitsuokella spp.	3	**Total**	**1574**
Whole Milk			
Genus	**Sequences**	**Genus**	**Sequences**
Anoxybacillus spp.	282	*Hallella spp.*	5
Bacillus spp.	9	*Lactococcus spp.*	89
Bacteroides spp.	38	*Microlunatus spp.*	7
Bifidobacterium spp.	2	*Pantoea spp.*	10
Caloramator spp.	2	*Prevotella spp.*	161
Chloroplast Unknown	2097	*Prosthecobacter spp.*	5
Chondromyces spp.	2	*Pseudomonas spp.*	340
Clostridium spp.	8	*Roseburia spp.*	2
Cylindrospermum spp.	5	*Ruminococcus spp.*	5
Desulfovibrio spp.	5	*Streptococcus spp.*	2
Escherichia spp.	11	*Streptomonospora spp.*	2
Friedmanniella spp.	2	*Succinivibrio spp.*	18
Fulvimonas spp.	3	*Tropheryma spp.*	16
Fusobacterium spp.	3	*Veillonella spp.*	2
		Total	**3133**

Table 12 Continued

Cheddar Cheese			
Genus	**Sequences**	**Genus**	**Sequences**
Lactococcus spp.	4605	*Streptococcus spp.*	2
		Total	**4607**
Spinach			
Genus	**Sequences**	**Genus**	**Sequences**
Pseudomonas spp.	32	*Oscillatoria spp.*	2
Pantoea spp.	16	*Tropheryma spp.*	17
Psychrobacter spp.	2	*Janthinobacterium spp.*	5
Cylindrospermum spp.	2018	*Chlorogloeopsis spp.*	4
		Total	**2096**

[α] Includes sequences rejected from final analysis because of an E-value above 10^{-80}.

CONTACT INFORMATION

Name:	Katherine Grace McElhany
Professional Address:	c/o Dr. Suresh Pillai Department of Poultry Science Texas A&M University College Station, TX 77843
Email Address:	kmcelhany@gmail.com
Education:	B.S. Biology, Texas A&M University, May 2008 Magna Cum Laude Undergraduate Research Scholar